

Rules and meta-rules in the framework of possibility theory and possibilistic logic¹

Didier Dubois¹, Henri Prade¹, Steven Schockaert²

¹Institut de Recherche en Informatique de Toulouse
University of Toulouse, 31062 Toulouse Cedex 09, France
{dubois | prade}@irit.fr

²Dept. of Applied Mathematics and Computer Science
Ghent University, Krijgslaan 281, 9000 Gent, Belgium
steven.schockaert@ugent.be

Running title : Rules and meta-rules in possibilistic logic

Corresponding Author : Didier Dubois (Fax : +33561556331).

Abstract:

The contribution of Lotfi Zadeh to the development of fuzzy logic, goes far beyond the introduction of the seminal concept of a fuzzy set and has multiple facets. This article, as a small tribute to the corpus of ideas, notions and results brought together over almost five decades by Zadeh, singles out and illustrates two of his most stimulating, thought-provoking and fruitful creations: fuzzy rules on the one hand, and possibility theory on the other hand. Indeed, the modeling of conditional statements of the form "if x is A then y is B " plays a crucial role in any attempt at formalizing human reasoning. Starting from the expression of different forms of fuzzy rules that have been identified in the setting of possibility theory, we study their counterparts in the extensions of possibilistic logic. A distinction between rules and meta-rules is especially emphasized, in the representational setting of possibility theory. It amounts to viewing rules as pieces of knowledge that contribute to the partial specification of a unique epistemic state, while meta-rules characterize constraints between specified epistemic states, as in possibilistic answer set programming.

Keywords:

Rule; meta-rule; fuzzy rule; possibility theory; possibilistic logic.

1 General introduction

Among the tremendous amount of contributions and ideas proposed and developed by Lotfi Zadeh over 45 years in the framework of fuzzy logic, we would like to pinpoint three very important pieces that have been especially influential for our own works: the notion of fuzzy rules [1], the new setting of possibility theory for modeling epistemic uncertainty [2], and its application to knowledge representation and approximate reasoning [3, 4]. After being controversial for some time the idea of fuzzy sets has finally been worldwide acknowledged as a simple and powerful tool in information modeling. It has been extremely successful in automatic control applications, where fuzzy rules are used as a way to easily implement interpolation mechanisms and as universal approximators of control laws.

Zadeh has also been continuously interested in artificial intelligence (AI) along his career, and this interest can be traced back as early as 1950 [5]! However, Zadeh's contribution to knowledge representation and approximate reasoning has remained less widely known, in spite of (or because) of its novelty when proposed in the late seventies. Indeed, the idea of representing pieces of information by means of possibility distributions, and of computing the results of inferences by the combination and projection of these distributions (in complete agreement with possibility theory) may be regarded as a breakthrough which anticipated many other works in AI ranging from constraint satisfaction problems to uncertainty networks. Even if Zadeh's proposal was mainly

¹ This paper is heartily dedicated to Lotfi Zadeh at the occasion of his 90th birthday, commemorating 6 decades of scientific contributions.

stated at the semantic level, it is sufficiently rich to include, as particular cases, possibilistic logic [6, 7, 8] which corresponds to the encoding of uncertain crisp propositions, as well as the extreme case of completely informed situations described by fuzzy propositions in a multiple-valued logic [9].

Many of our contributions to AI, e.g. [6, 10, 11, 12, 13], have thus not just been influenced by Zadeh's pioneering works, but in fact are deeply rooted in the setting of possibility theory that we have further developed and/or applied in new problems. Moreover, we are clearly indebted to Zadeh's contributions not only for their contents, but also for providing numerous examples of non-conventional views for approaching problems. In the following, we focus the discussions on fuzzy rules, a topic that may be found emblematic of Zadeh's proposals, showing that some fresh meat can be still offered on this old topic, in relation with present concerns in logic programming.

2 Fuzzy rules

Conditional statements play a very important role in the expression of knowledge. Fuzzy conditional statements viewed as expressions of the form "if A then B" where A and B have fuzzy meanings have been considered rather early in the development of fuzzy logic. They were introduced in a famous seminal paper [1] by Zadeh, which is at the basis of the blossoming of fuzzy rule-based control and decision systems. Interestingly enough, Zadeh's view of fuzzy rules [14, 15] has its roots in the idea of a fuzzy graph as a way of describing (input, output) pairs in a fuzzily described system [16, 17]. Such a fuzzy graph, associated with a fuzzy mapping, is made of fuzzy points (A, B) that can be read as a fuzzy rule "if A then B". Fuzzy points lead to the conjunctive view of fuzzy rules in terms of fuzzy Cartesian products $A \times B$, put at work by Mamdani and Assilian [18]'s fuzzy controllers, and their numerous followers. However, it is worth noticing that fuzzy conditional statements in [1] are discussed in terms of a material implication linking A and B. But the conjunctive and implicative views are there "reconciliated" by understanding $A \rightarrow B$ (where A and B are defined on domains U and V respectively and + denotes a disjunction), as $A \times B + \neg A \times V$, i.e. the model of the rule is on the one hand equivalent to a pair of fuzzy points, but on the other hand, it gives birth to the multiple-valued implication $\max(\min(a, b), 1 - a)$, or to Dienes' implication $\max(b, 1 - a)$ when replacing $A \times B$ by $U \times B$ in the previous expression, which leads to $U \times B + \neg A \times V$ (i.e., either one has B or not A) [1].

This ambiguous nature of if-then rules, whether fuzzy or not, can be related to what is often termed as "paradoxes of implication", noticed for a long time. Indeed, in propositional logic, material implication $A \rightarrow B$ imposes joint conditions on the truth of A and B, without really conveying any sense of necessity or relevance [19] in the so-established link between A and B (especially when A is false). This has led to many theoretical developments that address this problem in different ways: modal logics, conditional logics, fuzzy logics, logic programming, to name only the most visible ones.

Artificial intelligence has contributed to the emergence of the idea of a state of knowledge (or *epistemic state*) as a representation of what an agent could know about the world in given circumstances, see e.g. [20]. The framework of possibility theory [2] is particularly well-suited for a graded representation of such incomplete states of information, where it is generally useful to rank-order the possible values of n-tuples of variables describing the state of the world according to their level of plausibility. Indeed, the idea of a possibility distribution as an elastic constraint restricting

the possible values of a variable, introduced by Zadeh [2], is a key tool for representing incomplete states of information and fuzzy granules of knowledge. Zadeh [21] has further investigated the idea of using constraints as an information representation tool.

For years, researchers in possibility theory and fuzzy set theory have been interested in the proper modeling of if-then rules. This has led to the identification of different types of rules that involve the gradual nature of the properties and / or the uncertainty of the conclusions. As a result, a synthesis has finally emerged [22], which, within the framework of possibility theory, contrasts two types of modeling for rules, either at an “object” level or at a “meta level”. In this article, we reconsider this distinction, showing its interest and its counterpart in possibilistic logic [6, 7, 8].

The paper is structured as follows. First, we provide the necessary background about possibility theory, highlighting the four sets of functions associated with a possibility distribution and their role in the specification of these distributions. Then, we discuss the modeling of if-then rules in the framework of possibility theory, recalling the origin of the “conjunctive” and the “implicative” models, and stresses the distinction between the “object” and “meta” points of view in the modeling of rules, which is a key issue for the understanding of the different approaches. The next section explains how possibilistic logic can be generalized so as to take into account the two points of view, and we discuss the consequences for inference. Finally, we conclude the article by showing the importance of the distinction between rules and meta-rules for a better understanding of possibilistic logic programming.

3 Basics of possibility theory

This section provides a short background on possibility theory, starting with the basic notion of a possibility distribution [2], recalling how events can be estimated on this basis by means of different “measures”, and discussing how a possibility distribution can be indirectly specified through these measures.

3.1 Possibility distribution and associated measures

Given a universe of discourse U , a possibility distribution [2] π is a function from U to $[0, 1]$ that restricts the more or less possible values for a “state” whose exact value is ill-known. This state may be a quantity x with values in U , or the true, unique state of an incompletely described world. Then $\pi(u) = 0$ expresses that the value u is impossible for this ill-known quantity (or state), and u is all the more possible, as a value of this quantity, as $\pi(u)$ is large and close to 1.

Given a subset A in U , four set functions can be defined in the framework of possibility theory, namely $\forall A \subseteq U$:

- a (weak) possibility measure [2] :

$$\Pi(A) = \sup_{u \in A} \pi(u);$$

it estimates the compatibility / consistency of having A true (i.e. one of the elements of A is the true value) with the state of knowledge represented by π ; this may be seen as a measure of *potential* possibility;

- a dual measure of (strong) necessity [23] :

$$N(A) = 1 - \Pi(A^c) = \inf_{u \notin A} 1 - \pi(u);$$

it echoes the impossibility of the values outside A (A^c denotes the complement of A), or equivalently the inference of A from π and thus reflects an *actual* necessity;

- a (strong) measure of possibility [24] :

$$\Delta(A) = \inf_{u \in A} \pi(u);$$

it expresses to what extent *all* values in A have a minimal (guaranteed) level of possibility; this may be seen as a measure of *actual* possibility;

- a dual measure of (weak) necessity [24] :

$$\nabla(A) = 1 - \Delta(A^c) = \sup_{u \notin A} 1 - \pi(u);$$

it expresses to what extent there exists at least one impossible value outside A ; as such, this is a measure of *potential* necessity.

These measures are related together by the weak consistency constraint:

$$\forall A \subseteq U, \max(\Delta(A), \nabla(A)) \leq \min(\Pi(A), N(A));$$

it holds if $\exists u^* \in U \pi(u^*) = 1$ and $\exists u_* \in U \pi(u_*) = 0$ (in other words, if the possibility distribution is consistent and non-trivial).

3.2 Specification of a possibility distribution

A possibility distribution π is often only partially specified through constraints stating upper and lower bounds, respectively under the form $\pi \leq G$ and $F \leq \pi$, where F and G can be seen as the characteristic functions of fuzzy sets.² Let us notice that Π and N can still be used in a non-trivial way when one only knows that $\pi \leq G$ (since Π and N are increasing in the broad sense with respect to set inclusion), while Δ and ∇ are still appropriate if it is only known that $F \leq \pi$ (indeed Δ and ∇ are decreasing).

Clearly, in case of multiple restrictions G_i we get $\pi \leq \min_i G_i$, i.e. any information of this kind can only decrease the degrees of potential possibility, while we have $\max_j F_j \leq \pi$, which, on the contrary leads to an increase of the degrees of actual possibility.

In the following, we only consider possibility distributions with a finite number of possibility degrees for the sake of simplicity:

$$a_1 = 1 > a_2 \dots > a_{n+1} = 0, \text{ where } n \geq 1.$$

Moreover, these degrees are supposed to be chosen such that $1 - a_j = a_{n-j+2}$, which expresses that complementation is an internal operation on this scale.

Let $a_i(\pi) = \{u \mid \pi(u) \geq a_i\}$ be the cut of level a_i of π . Then, the distribution π can be decomposed under the form of the weighted disjunction of its n level cuts:

$$\pi(u) = \max_{i=1, \dots, n} \min(a_i(\pi)(u), a_i)$$

as well as a weighted conjunction:

$$\pi(u) = \min_{j=1, \dots, n} \max(a_j(\pi)(u), a_{j+1}).$$

² A subset and its characteristic function are denoted in the same way.

Thus, a possibility distribution π can be specified in terms of a finite number of level cuts, according to the following expressions:

$$\begin{aligned} \forall i = 1, n, \forall u, \min(F_i(u), a_i) \leq \pi(u) &\Leftrightarrow \Delta(F_i) \geq a_i; \\ \forall j = 1, n, \forall u, \pi(u) \leq \max(G_j(u), a_{j+1}) &\Leftrightarrow N(G_j) \geq 1 - a_{j+1}. \end{aligned}$$

When the fuzzy statement "x is G" is understood as the fact that the possibility distribution π_x representing this piece of information should satisfy the constraint $\pi_x \leq G$, choosing a particular distribution π_x such that $\pi_x < G$ would be arbitrarily too precise. This leads to taking π_x equal to the characteristic function of G. This is the *minimal specificity* principle [23]. Similarly, in the case of a constraint of the form $F \leq \pi_x$, one is led to take π_x equal to F (at least if $F \subseteq G$), thus applying a *maximal specificity* principle.

If we now consider negative statements such as "x is not F", they can be understood in two different ways:

- either as "x is (not F)", i.e., $\pi_x \leq F^c$, with $F^c(u) = 1 - F(u)$
- or as "x (is not) F" (that is "x is F" is not true), leading to : $\text{not}(\pi_x \leq F) \Leftrightarrow \exists u, \pi_x(u) > F(u)$.

These two interpretations coincide only if the value of x is precisely known, i.e., if π_x is equal to 1 for a unique value (and is 0 elsewhere).

4 Representations of if-then rules in possibility theory

This section first recalls the conjunctive and implicative modelings of a rule [25], prior to distinguishing between the "object" and "meta" points of view.

4.1 A first dichotomy : conjunction v.s. implication

Let us consider a rule of the form "if $x \in A$ then $y \in B$ ", where x and y are two quantities taking their values respectively on the domains U and V, where A and B are subsets of U and V. Let us first examine the case where A and B are ordinary subsets.

A and B non fuzzy. The partial description offered by a rule may be understood in two different ways : either one insists on what is positively asserted (namely that, when $x \in A$, any value in B is eligible for y), or in an implicitly negative way (the values outside B are excluded when $x \in A$):

- $(x, y) \in A \times B$ is guaranteed to be possible,
- $(x, y) \in A \times B^c$ is guaranteed to be impossible.

Since nothing is said for the situations where $x \in A^c$, if $\pi_{(x,y)}$ denotes the possibility distribution restricting the possible values of the pair (x, y), thus we have respectively :

- for the conjunction-based model (positive reading):

$$A \times B \leq \pi_{(x,y)} \quad (\Delta 1)$$

which leads to the conjunctive representation:

$$\pi_{*(x,y)}(u,v) = 1 \text{ if } (u, v) \in A \times B,$$

and $\pi_{*(x,y)}(u,v) = 0$ otherwise (since nothing is said for the other values).

As it may be expected we have, for all $\pi \geq \pi_*$:

$$\Delta(A \times B) = 1 \quad (\Delta 2)$$

Since, when $A(u) = 1$ we have $\pi_{*(x,y)}(u,v) = 1$ if $v \in B$, $(\Delta 1)$ is equivalent to:

$$A(u) \geq \inf_{v \in B} \pi_{*(x,y)}(u,v) = \Delta_*(\{u\} \times B) \quad (\Delta 3)$$

i.e., if A is true (for $x = u$), B is indeed guaranteed to be possible. By symmetry, $(\Delta 1)$ is equivalent to

$$B(v) \leq \inf_{u \in A} \pi_{*(x,y)}(u,v) = \inf_{u \in U} \max(1 - A(u), \pi_{*(x,y)}(u,v)) \quad (\Delta 4)$$

which corresponds to the proper inference mode from what is guaranteed to be possible (see [26] for details).

- for the implication-based model (negative reading), we have:

$$\pi_{(x,y)} \leq (A \times B^c)^c = A^c + B \quad (N1)$$

which expresses that $\pi_{*(x,y)}(u,v) = 0$ if $(u,v) \in A \times B^c$,

and $\pi_{*(x,y)}(u,v) = 1$ otherwise (since the other values are not restricted).

As expected we have, for all $\pi \geq \pi_*$

$$N(A^c + B) = 1 \quad (N2)$$

Besides $(N1)$ is equivalent to

$$A(u) \geq \inf_{v \notin B} 1 - \pi_{*(x,y)}(u,v) = N^*(\{u\} \times B) \quad (N3)$$

i.e., if A is true (for $x = u$), $y \in B$ is certain.

Moreover, $(N1)$ is equivalent to

$$\sup_{u \in U} \min(A(u), \pi_{*(x,y)}(u,v)) \leq B(v), \quad (N4)$$

which is the counterpart of the classical inference: from a state of knowledge including the (implicative) representation of the rule, and the one of A is true, one indeed deduces that B is true. The inference machinery at work in $(N4)$ is Zadeh's combination/projection principle [4], or more simply here what is also called the compositional rule of inference, together with the entailment principle, which are at the basis of his theory of approximate reasoning.

We thus have two possible views of a rule "if A then B" leading to a representation of its epistemic contents that associates A and B under the form of either an implicative relation π^* which is minimally specific, or as a conjunctive relation π_* which is maximally specific. It corresponds to a bipolar view of a rule [27], the positive side focusing on examples characterized by the conjunction

of antecedent and consequent and the negative side focusing on counter-examples characterized by the complement of the material implication.

Remark. When A and B are fuzzy, $(\Delta 1)$ and $(N1)$ can be extended in different ways according to the choice of the conjunction \times (or of the implication for $N1$), leading to six types of rules [22], where the level cuts of A and B are associated in different ways. Namely, depending on whether the implication $A \rightarrow B$ is modeled by a strong implication (as Dienes'), a residuated implication (as Gödel's), or its contrapositive, we get a certainty rule, a gradual rule, or an impossibility rule. Depending on whether, for modeling \times in $A \times B$, we use a triangular norm (for example the minimum), or the right, or the left adjoint of a residuated implication, we get a (guaranteed) possibility rule, or two types of "anti-gradual" rules. More details in [22, 25].

4.2 A second dichotomy : rules vs. meta-rules

As already said, in the previous approach, in the case where the possibility distributions are viewed as restrictions on possible values [2], the inference is governed by the combination / projection principle [4] as is clear in $(N4)$. However, some works have considered the inference process directly at the meta level by directly specifying π_y as soon as the condition part of the rule is satisfied. This idea has been studied in particular by Esteva et al. [28] (see also [22]); it also underlies the "compatibility-modification inference" proposed by Cross and Sudkamp [29].

Apart from the bipolar relational view of an if-then rule just recalled, we may indeed see the rule "if x is A then y is B " as the specification of a link between constraints on the possibility distributions π_x and π_y which describe states of knowledge about the quantities x and y respectively [22]. Rules of this latter type are called meta-rules, as they encode relationships between different epistemic states, at the meta-level, rather than relationships between possible values, at the object level. Each of the two components of the rule may be then interpreted as, for instance, $\pi_x \leq A$ and $\pi_y \leq B$ (by modeling 'x is A' and 'y is B' by restrictions on the possible values of x and y). This leads to see the rule as a crisp 'production' rule of the form

$$\text{if } \pi_x \leq A \text{ then } \pi_y \leq B \quad (NN1)$$

which can also be interpreted in terms of necessity measures, as:

$$\text{if } N_x(A) = 1 \text{ then } N_y(B) = 1 \quad (NN2)$$

i.e., "if A is certain then B is certain", where N_x and N_y are necessity measures associated to the possibility distributions π_x and π_y respectively.

Obviously, the statement "if x is A then y is B " understood in this way is logically equivalent in classical logic to "not(x is A) or y is B ", i.e.,

$$\text{not}(\pi_x \leq A) \text{ or } \pi_y \leq B.$$

A second type of meta-rules results from the observation that, apart from $\text{not}(\pi_x \leq A)$ understood as $\exists u \pi_x(u) > A(u)$, there exists another interpretation of negation that corresponds to $\pi_x \leq 1 - A$.

Then if we consider the statement "(x is A^c) or y is B", i.e., $\pi_x \leq 1 - A$ or $\pi_y \leq B$, one sees that it may be rewritten under the form

$$\text{if not}(\pi_x \leq A^c) \text{ then } \pi_y \leq B$$

The latter, when A and B are ordinary subsets, may be read in terms of a Boolean possibility measure, since then $\text{not}(\pi_x \leq A^c) \Leftrightarrow \text{not}(N_x(A^c) = 1) \Leftrightarrow N_x(A^c) = 0 \Leftrightarrow \Pi_x(A) = 1$.

This gives birth to a new kind of rule of the form

$$\text{if } \Pi_x(A) = 1 \text{ then } N_y(B) = 1 \quad (\Pi N1)$$

i.e., "if A is possible then B is certain".

Note that since $\pi_{x,y} \leq \pi_x \leq 1 - A$ or $\pi_{x,y} \leq \pi_y \leq B^3$ entails $\pi_{x,y} \leq \max(1 - A, B)$, ($\Pi N1$) is stronger than ($N1$), and since $\pi_x \leq 1 - A$ entails $\text{not}(\pi_x \leq A)$, assuming that π_x is normalised, ($\Pi N1$) is also stronger than ($NN1$).

In a similar way, two other 'production' rules can be defined

- under the form

$$\text{if } \pi_x \geq A \text{ then } \pi_y \geq B \quad (\Delta\Delta1)$$

corresponding to

$$\text{if } \Delta_x(A) = 1 \text{ then } \Delta_y(B) = 1 \quad (\Delta\Delta2)$$

i.e., if A is actually possible, B is too;

- and under the form

$$\text{if } \nabla_x(A) = 1 \text{ then } \Delta_y(B) = 1 \quad (\nabla\Delta1)$$

i.e., "if some values outside A are impossible then B is guaranteed to be possible".

5 Rules and meta-rules in possibilistic logic

We now study the counterpart of these rules, in a graded version, in the setting of possibilistic logic. We start with a brief reminder on possibilistic logic, a weighted extension of classical logic, where the weights are handled using possibility theory.

5.1 Possibilistic logic: a short account

A formula in standard possibilistic logic [6, 7] is a pair (p, a) where p is a classical logic proposition, and a its level of certainty. It is semantically interpreted under the form of the constraint $N(p) \geq a$ ($\Leftrightarrow \Pi(\neg p) \leq 1 - a$), and is associated to the possibility distribution

$$\pi_{(p,a)}(s) = \max(\langle p \rangle(s), 1 - a)$$

where $\langle p \rangle(s) = 1$ if s is a model of p (i.e. an interpretation that makes p true) and $\langle p \rangle(s) = 0$ otherwise (the interpretations that make p false are possible (at most) at degree $1 - a$).

³ Indeed, by definition, $\pi_x(u) = \sup_v \pi_{x,y}(u, v) \geq \pi_{x,y}(u, v)$, and $\pi_y(v) = \sup_u \pi_{x,y}(u, v) \geq \pi_{x,y}(u, v)$.

Inference is based on the cut rule

$$(\neg p \vee q, a); (p \vee r, b) \models (q \vee r, \min(a, b)).$$

Note that the modus ponens instance of this rule, namely $(\neg p \vee q, a); (p, b) \models (q, \min(a, b))$, is in complete agreement with the compositional rule of inference. Namely, it reads semantically

$$\min(\pi_{(p, b)}(s), \pi_{(\neg p \vee q, a)}(s)) \leq \pi_{(q, \min(a, b))}(s)$$

i.e., $\min(\max(\langle p \rangle(u), 1 - b), \max(1 - \langle p \rangle(u), \langle q \rangle(u), 1 - a)) \leq \max(\langle q \rangle(u), 1 - \min(a, b))$,

which is an instance of

$$\sup_u \min(\pi_x(u), \pi_{x,y}(u, v)) \leq \pi_y(v)$$

with π_x representing « x is $\langle p \rangle$ is b -certain », $\pi_{x,y}$ representing « (if x is $\langle p \rangle$ then y is $\langle q \rangle$) is a -certain », and π_y representing « y is $\langle q \rangle$ is $\min(a, b)$ -certain ». In the same spirit, in a research note, Zadeh [30] has pointed out the parallel between the compositional rule of inference and a counterpart of Prolog where expressions are associated with certainty weights equal to 1. See also [9] for related discussions.

Besides, there also exists a logic in terms of guaranteed (actual) possibility [8] with formulas of the form $[p, a]$, where p is a proposition, and a is its guaranteed possibility level. This corresponds semantically to the constraint $\Delta(p) \geq a$ ($\Leftrightarrow \nabla(\neg p) \leq 1 - a$), and is associated with the possibility distribution

$$\pi_{[p, a]}(u) = \min(\langle p \rangle(u), a).$$

The corresponding inference is based on the rule

$$[\neg p \wedge q, a]; [p \wedge r, b] \models [q \wedge r, \min(a, b)].$$

In order to generalize (NN2) for example, one should express that $N(p) \geq a \Rightarrow N(q) \geq a$. But note that even by enforcing $a < b$ in the pair of formulas $(p, a), (q, b)$, possibilistic logic does not enable us to express the above implication. One may also enforce the condition $N(q) \geq N(p)$, which would be even stronger than the above implication \Rightarrow . This latter type of information has already been considered in [31]. Similarly $\{[p, a], [q, b]\}$, with $a < b$, does not express that $\Delta(p) \geq a \Rightarrow \Delta(q) \geq b$. We are going to see how such relationships can be expressed in a generalized possibilistic logic.

5.2 Generalized possibilistic logic

We may now consider the counterpart of rules such as (NN2). Formulas in possibilistic logic may be linked by only one connective, the conjunction. Indeed the possibilistic logic base $\{(p, a), (q, b)\}$ is equivalent to the conjunction $(p, a) \wedge (q, b)$, which allows us to work in a clausal form. This is because $N(p) \geq a \wedge N(q) \geq b$ is semantically equivalent to the possibility distribution

$$\pi(u) = \min(\max(\langle p \rangle(u), 1 - a), \max(\langle q \rangle(u), 1 - b)).$$

in standard possibilistic logic. The formula $(p, a) \vee (q, b)$, which semantically corresponds to a set of two possibility distributions $\pi_{(p, a)}$ and $\pi_{(q, b)}$, should be understood at a meta-level with respect to one of the two elementary formulas that constitute it. This expresses the disjunctive constraint

$$N(p) \geq a \text{ or } N(q) \geq b.$$

In the same way, $\neg(p, a)$ expresses that it is false that $N(p) \geq a$ and thus $N(p) < a$. One may thus apply classical logic inference at the meta level to such formulas, as already suggested in [32].

Thus if we have the knowledge base $K = \{\neg(p, a) \vee (q, a), (p, a)\}$ one can deduce (q, a) , which is in accordance with the intuition that the epistemic state (q, a) can be produced, once the epistemic state (p, a) is established, given the meta-rule $\neg(p, a) \vee (q, a)$.

It is interesting to observe that while formula $(\neg p \vee q, a)$, enables us to deduce both (q, a) if one has (p, a) , and $(\neg p, a)$ if one has $(\neg q, a)$, the meta-formula $\neg(p, a) \vee (q, a)$ also enables us to get (q, a) from (p, a) , but does not enable us to deduce $(\neg p, a)$ in the presence of $(\neg q, a)$. Indeed, $(\neg q, a)$ expresses that $N(\neg q) \geq a$, which entails $N(q) = 0$ if $a > 0$ (since $\min(N(q), N(\neg q)) = 0$ holds). Knowing $N(q) = 0$, along with the rule “ $N(p) < a$ or $N(q) \geq a > 0$ ” entails $N(p) < a$, which differs from the stronger conclusion $(\neg p, a)$, i.e., $N(\neg p) \geq a$, that may be obtained from $(\neg p \vee q, a)$ and $(\neg q, a)$. Moreover $N(p) < a$, i.e., $\neg(p, a)$ is not a possibilistic logic formula at the object level. This shows the deep difference between $(\neg p \vee q, a)$ and $\neg(p, a) \vee (q, a)$, or between $\neg(p, a)$ and $(\neg p, a)$.

In the same way, the counterpart of the meta-rule ($\Pi N1$) writes

$$\text{if } \Pi(p) \geq a \text{ then } N(q) \geq a,$$

i.e., $\neg(p, a) \vee (q, a)$, where a formula of the form $\langle p, a \rangle$ encodes the constraint $\Pi(p) \geq a$. A hybrid resolution rule [7,8] enables us to reason from such clauses:

$$(\neg p \vee q, a); \langle p \vee r, b \rangle \mid = \langle q \vee r, b \rangle \text{ if } b > 1 - a.$$

If one wants to express, as in logic programming with negation as failure, that

“ r is certain provided that p is certain and that one cannot establish q ”

it leads to write is as

$$\text{if } N(p) \geq a \text{ and } \Pi(\neg q) \geq b \text{ then } N(r) \geq a,$$

which corresponds to formula $\neg(p, a) \vee \neg(\neg q, b) \vee (r, a)$.

Automated reasoning in this setting presupposes to deal with three kinds of information:

- “facts” that are more or less certain, encoded by standard possibilistic logic formulas of the form (p, a) where p is a clause and $a > 0$.
- possibilistic formulas (q, a) that are “currently impossible to establish” for any $a > 0$, which may be expressed under the form $\langle \neg q, 1 \rangle$, since $\Pi(\neg q) = 1$ is equivalent to $N(q) = 0$. Such

kinds of statements, that result from a possible lack of information, may also be propagated by means of the hybrid resolution rule, once they are acknowledged and expressed. It is clear that the arrival of new pieces of information which may enable us to establish new conclusions may also lead to delete formulas of the type $\langle \neg q, 1 \rangle$ that were previously accepted.

- meta-rules of the form $\neg(p, a) \vee (r, a)$, or more generally $\neg(p, a) \vee \neg\langle \neg q, 1 \rangle \vee (r, a)$, which enable to produce from the two first types of information, *positive* pieces of information of the form (r, a) .

This view raises questions on the links between such a generalized possibilistic logic and on the one hand, possibilistic logic programming [33, 34], and on the other hand, the logical manipulation of epistemic states in modal logic style [35], or with maybe the explicitation of some forms of “(non) awareness” [36]. The following and last section discusses the possibilistic logic interpretation of classical Answer Set Programming (ASP).

5.3 Answer sets in generalized possibilistic logic

The aim of this section is to summarize how the semantics of answer set programming can be described in the generalized possibilistic logic that was introduced above. Recall that an answer set program is a set of rules of the following form:

$$r \leftarrow p_1, \dots, p_n, \text{not } q_1, \dots, \text{not } q_m \quad (\text{R1})$$

with the intuitive meaning that whenever p_1, \dots, p_n can be established, we may establish r , unless one of q_1, \dots, q_m can be established. Answer sets are consistent sets of literals that can be derived from a set of such rules using non-deterministic forward chaining. The non-determinism results from the fact that when applying a rule such as (R1), we need to make the assumption that neither of q_1, \dots, q_m will be established during the forward chaining procedure. An answer-set of a program P is represented by the set of positive literals A in this model (subset of the so-called Herbrand universe H , i.e. the set of all atoms appearing in P). A model of a rule like R1, is then a subset $A \subseteq H$ such that A contains r , all p_i 's and no q_j 's.

The rule (R1) corresponds to the following formula in possibilistic logic:

$$(r, 1) \vee \neg(p_1, 1) \vee \dots \vee \neg(p_n, 1) \vee \neg\langle \neg q_1, 1 \rangle \vee \dots \vee \neg\langle \neg q_m, 1 \rangle \quad (\text{R2})$$

In particular, note that a fact of the form $\langle r \leftarrow \rangle$ is translated to the formula $(r, 1)$. For programs P without negation-as-failure (i.e. $m=0$), there is no non-determinism. In such a case, P has exactly one answer set A , which is its unique minimal model with respect to inclusion (when interpreting the rules in classical logic, using material implication). In such a case, the proposed translation into possibilistic logic fully captures the answer set semantics, as is revealed by the following proposition.

Proposition 1. Let P be an answer set program without negation, and let K_p be its translation in generalized possibilistic logic, as described above. Let A be the unique answer set of P . It holds that $p \in A$ iff $K_p \models (p, 1)$.

Proof (sketch). The proof follows straightforwardly from the fact that applying the resolution rule to the formulas in K_p corresponds to applying forward chaining.

QED

The presence of negation-as-failure introduces non-determinism, resulting in the fact that a program P may have several answer sets, or none at all. Given a potential answer set A (representing an interpretation), checking if it is a model of P comes down to first computing the Gelfond-Lifschitz reduct of P with respect to A , denoted P^A , and obtained as follows: delete from P all rules trivially satisfied by A , that is, rules such that $q_j \in A$. Then P^A is the set of rules obtained by deleting the negative part of the body of all remaining rules. The model A is an answer set of P if A is the unique minimal model of P^A .

To encode the answer set semantics in possibilistic logic, we will need to consider two sets of formulas: the set K_p as before and a set of assumptions. Indeed, for programs with negation-as-failure, we face the problem that the possibilistic logic base K_p cannot semantically entail any formula of the form $\langle \neg q, 1 \rangle$ with a negative literal, hence the only inferences that can be made are based on the translation of rules in which no negation-as-failure occurs. We can now add assumptions of the form $\langle \neg q, 1 \rangle$ to K_p , meaning that it is consistent to assume $\neg q$ (or more precisely, inconsistent to derive q with certainty). The non-determinism of the forward chaining procedure is thus translated in the choice of which assumptions of the form $\langle \neg q, 1 \rangle$ to consider. Specifically, for B a set of atoms, we define $M_B = \{ \langle \neg q, 1 \rangle \mid q \in B \}$. Then we have the following result.

Proposition 2. Let P be an answer set program. It holds that A is an answer set of P iff $K_p \cup M_{HVA}$ is consistent and $K_p \cup M_{HVA} \models (p, 1)$ for all $p \in A$.

Proof (sketch). The fact that for an answer set A of P , it holds that $K_p \cup M_{HVA}$ is consistent, follows easily from the fact that $K_{p^A} \cup M_{HVA}$ is consistent, which in turn follows from Proposition 1 (with P^A the Gelfond-Lifschitz reduct of P). Similarly, $K_p \cup M_{HVA} \models (p, 1)$ follows from the observation that $K_{p^A} \cup M_{HVA} \models (p, 1)$ due to Proposition 1. Conversely, if $K_p \cup M_{HVA}$ is consistent and $K_p \cup M_{HVA} \models (p, 1)$ for all $p \in A$, we need to show that A is the answer set of the reduct P^A . By observing that the only inconsistencies that may exist between disjuncts of K_{p^A} and formulas of M_{HVA} are between formulas of the form $\langle q, 1 \rangle$ and $\langle \neg q, 1 \rangle$, it is easy to show that the answer set A^* of P^A should be such that $A \subseteq A^*$, using Proposition 1 and the assumption that $K_p \cup M_{HVA} \models (p, 1)$ for all $p \in A$. From the assumption that $K_p \cup M_{HVA}$ is consistent, we moreover find that $A^* \subseteq A$.

QED

Note that the condition that $K_p \cup M_{HVA} \models (p, 1)$ can alternatively be written as " $K_p \cup M_{HVA \setminus \{p\}}$ is inconsistent for all $p \in A$ ". In other words, answer sets correspond to maximal sets of assumptions of the form $\langle \neg q, 1 \rangle$ consistent with K_p .

In the weighted case, note that, for any possibilistic logic base K we either have $K \models \langle q, a \rangle$ or $K \models (\neg q, b)$ for some $b > 1 - a$, as $\langle q, a \rangle$ corresponds to the constraint $\Pi(q) \geq a$ while $(\neg q, b)$ corresponds to $\Pi(q) \leq 1 - b$. This means that, under the minimal specificity semantics, adding formulas of the form $\langle q, a \rangle$ is either redundant at the semantic level or makes the resulting possibilistic logic base inconsistent. This suggests to read $\langle q, a \rangle$ as "it is consistent to assume that q is possible (to degree a)". Hence, we could look at formulas of the form (p, a) as expressing knowledge, and at formulas of the form $\langle q, a \rangle$ as expressing constraints on what may be derived. Now, assume that K is a standard possibilistic logic base and that $K \cup \{ \langle q, a \rangle \}$ is consistent, then K and $K \cup \{ \langle q, a \rangle \}$ are semantically equivalent in the sense that they induce the

same possibility distribution. When we consider an expression of the form $(p, a) \vee \neg\langle q, a \rangle$, however, the picture changes. In particular, the following two bases will typically not be equivalent anymore (provided that $\langle q, a \rangle$ is not a standard possibilistic consequence of K):

$$\begin{aligned} K' &= K \cup \{(p, a) \vee \neg\langle q, a \rangle\} \\ K'' &= K \cup \{\langle q, a \rangle\} \cup \{(p, a) \vee \neg\langle q, a \rangle\} \end{aligned}$$

In particular, note that $\{\langle q, a \rangle\} \cup \{(p, a) \vee \neg\langle q, a \rangle\} \models (p, a)$. If $(p, a) \cup K \models (\neg q, b)$ for some weight $b > 1 - a$, we thus find that K'' is inconsistent, whereas K' may not be.

This apparent paradox can be solved using a semantics in terms of families of possibility distributions. Define $\pi \models (p, a)$ if and only if $N(p) \geq a$, and $\pi \models \langle p, a \rangle$ if and only if $\Pi(p) \geq a$. Rather than identifying a possibilistic knowledge base with a single possibility distribution, being the least specific possibility distribution satisfying the imposed constraints, we may also identify a possibilistic knowledge base with the set of all possibility distributions that satisfy these constraints. In other words, we use meta-models [32, 35] consisting of sets of possibility distributions satisfying possibilistic formulas. Under this extended semantics, it is clear that the set of meta-models of K and the one of $K \cup \{\langle q, a \rangle\}$ are generally different. In other words the above discussion suggests that:

- Generalizing possibilistic logic with more connectives and modalities requires moving from weighted rules to meta-rules at the syntactic level (where formulas relate belief states, rather than only restricting possible worlds), and moving from a fuzzy set of models (a single possibility distribution) to a set of fuzzy meta-models, at the semantic level.
- This seems to be the only way to go in order to capture and extend answer set programming in the setting of possibility theory.

6 Conclusion

The article has sketched an overview of the different forms of if-then rules that can be expressed in the framework of possibility theory, emphasizing the difference between the rules that contribute to specify an epistemic state, and the meta-rules that go from a partial epistemic state to another partial epistemic state. This distinction has enabled us to bridge a generalization of possibilistic logic with possibilistic answer set programming. It is also worth pointing out that the meta-rules that are based on guaranteed possibility, suggested in this article for the first time, are still to be studied for a better understanding of their potential interest. Besides, the idea of processing approximate reasoning at a symbolic level [37], in agreement with the idea of computing with words [38], might be revisited in terms of meta-rules.

Remarkably enough, the theory of approximate reasoning and possibility theory are not among Zadeh's most cited contributions nowadays. This is due to the tremendous success of fuzzy controllers and the subsequent association between fuzzy rules and neural networks, which popularized fuzzy systems in the area of numerical information processing. This trend, which in some sense questions the linguistic expert-driven stance of Zadeh's pioneering works, has also cast some suspicion on the relevance of symbolic artificial intelligence, whose approach was presented as being at odds with fuzzy logic at large. But, from the beginning, Mamdani used to see fuzzy control as an application of Artificial Intelligence. Moreover, Zadeh's approximate reasoning theory, totally neglected by the neuro-fuzzy tradition, is a fuzzy version of logic-based AI (possibilistic logic is a particular case of it), and basic axioms of possibility theory turn out to lie at

the core of a major approach to non-monotonic reasoning [10, 11, 12]. These facts demonstrate that the contributions of Zadeh pioneered important works quite outside the fuzzy community (whether their tenants admit it or not). This short paper suggests a new bridge between possibility theory and one of the most popular approaches in symbolic AI to-date, answer-set programming. This bridge is also related to pioneering texts of Zadeh concerning “possibility-qualification” of linguistic statements [3], thus indicating that after 45 years of fuzzy logic, some contributions by its founder that are currently often neglected have a good chance to be rediscovered.

References

- [1] Zadeh, L.A. “Outline of a new approach to the analysis of complex systems and decision processes”, *IEEE Trans. Systems, Man and Cybernetics*, **3**, pp 28-44 (1973).
- [2] Zadeh, L.A. “Fuzzy sets as a basis for a theory of possibility”, *Fuzzy Sets and Systems*, **1**, pp 3-28 (1978).
- [3] Zadeh, L.A. “PRUF - A meaning representation language for natural languages”, *Int. J. of Man-Machine Studies*, **10**, pp 395-460 (1978).
- [4] Zadeh, L.A. “A theory of approximate reasoning” In *Machine Intelligence*, **9**, J.E. Hayes, D. Michie, and L.I. Mikulich, eds., pp 149-194, Elsevier, Amsterdam, The Netherlands (1979).
- [5] Zadeh, L.A. “Thinking machines. A new field in electrical engineering”, *Columbia Engineering Quarterly*, **3**, pp 12-13 & 30-31 (1950).
- [6] Dubois, D., Lang, J. and Prade H. “Automated reasoning using possibilistic logic: semantics, belief revision and variable certainty weights”, *IEEE Trans. on Data and Knowledge Engineering*, **6** (1), pp 64-71 (1994).
- [7] Dubois, D., Lang, J. and Prade H. “Possibilistic logic”, In *Handbook of Logic in Artificial Intelligence and Logic Programming*, **3**, D.M. Gabbay, C.J. Hogger, J.A. Robinson, D. Nute, Eds., pp 439-513, Oxford University Press, UK (1994).
- [8] Dubois, D. and Prade, H. Possibilistic logic: a retrospective and prospective view. *Fuzzy Sets and Systems*, **144**, 3-23 (2004).
- [9] Dubois, D., Lang, J. and Prade H. “Fuzzy sets in approximate reasoning. Part 2: Logical approaches”, *Fuzzy Sets and Systems*, **40**, pp 203-244 (1991).
- [10] Dubois, D. and Prade, H. “Possibilistic logic, preferential models, non-monotonicity and related issues”, *Proc. of the Inter. Joint Conf. on Artificial Intelligence (IJCAI'91)*, pp 419-424, Sydney (1991).
- [11] Benferhat, S., Dubois, D. and Prade, H. “Nonmonotonic reasoning, conditional objects and possibility theory”, *Artificial Intelligence*, **92**, pp 259-276 (1997).
- [12] Benferhat, S., Dubois, D. and Prade, H. “Practical Handling of Exception-Tainted Rules and Independence Information in Possibilistic Logic”, *Applied Intelligence*, **9**, pp 101-127 (1998).
- [13] Schockaert, S. De Cock, M. and Kerre E. *Reasoning about Fuzzy Temporal and Spatial Information from the Web*. World Scientific, Singapore (2011).
- [14] Zadeh, L.A. “The calculus of fuzzy if-then rules”, *AI Expert*, **7** (3), pp 23-27 (1992).
- [15] Zadeh, L.A. “Fuzzy logic, neural networks, and soft computing”, *Communications of the ACM*, **37** (3), pp 77-84 (1994).
- [16] Zadeh, L.A. “Toward a theory of fuzzy systems”, In *Aspects of Networks and Systems Theory*, R. E. Kalman, N. De Claris, Eds., Holt, pp 469-490, Rinehart and Winston, New York (1971).
- [17] Zadeh, L.A. “On the analysis of large scale systems”, In *Systems Approaches and Environment Problems*, H. Gottinger, Ed., pp 23-37, Vandenhoeck and Ruprecht, Gottingen, Germany (1974).
- [18] Mamdani, E. H. and Assilian S. “An experiment in linguistic synthesis with a fuzzy logic controller”, *International Journal on Man-Machine Studies*, **7**, pp 1-13 (1975).
- [19] Anderson, A. R. and Belnap, N. D. *Entailment. The Logic of Relevance and Necessity*, vol. 1. Princeton University Press (1975); vol. 2, edited collection with the collaboration of J. M. Dunn, Princeton University Press (1992).
- [20] Gardenförs, P. *Knowledge in Flux. Modelling the Dynamics of Epistemic States*, MIT Press, Cambridge, Ma., (1988).
- [21] Zadeh, L.A. “Toward a generalized theory of uncertainty (GTU) — principal concepts and ideas”, *Computational Statistics and Data Analysis*, **51**, 15-46, 2006.
- [22] Bouchon-Meunier, B. Dubois, D. Godo, L. and Prade H. “Fuzzy Sets and Possibility Theory in Approximate and Plausible Reasoning”, In *Fuzzy Sets in Approximate Reasoning and Information Systems*, J. Bezdek, D. Dubois, H. Prade, Eds., pp 15-190, Kluwer Acad. Publ., Dordrecht, The Netherlands (1999).
- [23] Dubois, D. and Prade, H. *Possibility Theory*, Plenum Press, New-York (1988).

- [24] Dubois, D. and Prade, H. "Possibility theory: qualitative and quantitative aspects", In *Handbook of Defeasible Reasoning and Uncertainty Management Systems*, **1**, D.M. Gabbay, P. Smets, Eds., pp 169-226, Kluwer Acad. Publ., Dordrecht, The Netherlands (1998).
- [25] Dubois, D. and Prade, H. "What are fuzzy rules and how to use them", *Fuzzy Sets and Systems*, **84**, pp 169-185 (1996).
- [26] Dubois, D. and Prade, H. and Ughetto L. "A new perspective on reasoning with fuzzy rules", *Inter. J. of Intelligent Systems*, **18**, pp 541-567 (2003).
- [27] Dubois, D. and Prade, H. "Bipolar representations in reasoning, knowledge extraction and decision", *Proc. Inter. Conf. on Rough Sets and Current Trends in Computing (RSCTC 2006)*, S.Greco, Y. Hata, S. Hirano, M. Inuiguchi, Eds., Kobe, Japan, LNAI **4259**, pp 15-26, Springer, Berlin (2006).
- [28] Esteva, F. Garcia, P. and Godo L. On the semantics of fuzzy statements based on possibilistic constraints. *Proc. VII Congreso Español sobre Tecnologías y Lógica Fuzzy (ESTYLF'97)*, Tarragona, pp 21-27 (1997).
- [29] Cross, V. Sudkamp, T. "Patterns of fuzzy rule-based inference", *Int. J. of Approximate Reasoning*, **11**, pp 235-255 (1994).
- [30] Zadeh, L.A. "QSA/FL -- Qualitative systems analysis based on fuzzy logic", *Working Notes of the Stanford AAAI Symp. on AI and Limited Rationality*, M. Fehling, S. Russell, Eds., pp 111-114, Stanford, Ca., 1989.
- [31] Halpern J. "Defining relative likelihood in partially-ordered preferential structures" *J. of Artificial Intelligence Research*, **7**, pp 1-24, (1997).
- [32] Dubois, D. and Prade, H. "Toward multiple-agent extensions of possibilistic logic", *Proc. IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2007)*, pp 187-192, London (2007).
- [33] Bauters, K. Schockaert, S. De Cock, M. and Vermeir D. "Possibilistic answer set programming revisited", *Proc. 26th Conference on Uncertainty in Artificial Intelligence (UAI 2010)*, Catalina Island (2010).
- [34] Nicolas, P. Garcia, L. and Stéphan I. "Possibilistic stable models", *Proc. of the 19th International Joint Conference on Artificial Intelligence (IJCAI 2005)*, L. P. Kaelbling, A. Saffiotti, Eds., pp 248-253, Edinburgh (2005).
- [35] Banerjee, M. and Dubois D. "A simple modal logic for reasoning about revealed beliefs", *Proc. European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU 2009)*, Verona, C. Sossai, G. Chemello, Eds., Springer, LNAI **5590**, pp 805-816 (2009).
- [36] Prade H. "Handling (un)awareness and related issues in possibilistic logic: A preliminary discussion". *Proc. International Workshop on Non-Monotonic Reasoning (NMR 2006)*, pp 219-225, Lake District, UK, Clausthal University of Technology (2006).
- [37] Dubois, D. Foulloy, L. Galichet, S. and Prade, H. "Performing approximate reasoning with words?", In: *Computing with Words in Information/Intelligent Systems*, **1**, L. A. Zadeh, J. Kacprzyk, Eds., Springer-Verlag, pp 24-49 (1999).
- [38] Zadeh, L.A. "Fuzzy logic = computing with words", *IEEE Trans. on Fuzzy Systems*, **4**, pp 103-111 (1996).